INTERNATIONAL JOURNAL OF MANAGEMENT AND ORGANIZATIONAL RESEARCH

Forecasting Feeder-Level Outages with Hybrid Time-Series/ML Models: Accuracy, Explainability, and Maintenance Prioritization

Leeman Takunda Gunzo [1*], **Munashe Naphtali Mupa** [2], **Tendai Nemure** [3], **Japhet Dalokhule Muchenje** [4], **Nomagugu T Ndlovu** [5]

[1] University of Tennessee Knoxville, USA
[2] Hult International Business School, USA
[3] Yeshiva University, USA
[4] Suffolk University, USA
[5] Clarkson University, USA

* Corresponding Author: **Leeman Takunda Gunzo**

Article Info

ISSN (online): 2583-6641
Volume: 04
Issue: 05
September – October 2025
Received: 26-07-2025
Accepted: 28-08-2025
Published: 23-09-2025
Page No: 71-78

Abstract
Feeder-level outages are a real headache for utilities nowadays, as climate variability, fast vegetation growth, and aging infrastructure are making the network more vulnerable. The current article gives a discussion on popular statistical time-series models like ARIMA, ETS, and Prophet and compares them to tree-based ensembles and LSTM hybrids in predicting outages in the short term. It is a combination of historical ticket, weather radar, vegetation indices, and asset data. The condition evaluation methodology is a rolling-origin condition evaluation methodology whereby the performance is measured in terms of accuracy, calibration, error statistics, and reliability plots. Findings indicate the existence of the hybrid models by performing better than the classical models, particularly where the nonlinear interaction between vegetation and the weather is considered. The outcomes of the calibration test also show that the hybrids are more reliable when in a hybrid format. To bridge the gap between the model performance and application, the study comprises the combination of SHAP-based attribution to graphically highlight the risk drivers in a clear manner, such as vegetation density and adverse weather. Such interpretability gives more straightforward priorities in terms of maintenance and regulatory compliance. The results indicate that hybrid plus explainable pipelines have the potential to make outage predictions turn into a viable utility-level decision-support system.

DOI: https://doi.org/10.54660/IJMOR.2025.4.5.71-78

Keywords: Feeder-Level Outage Forecasting, Hybrid Time-Series Models, Machine Learning, Prediction Accuracy

Introduction

The stability of the reliability of the distribution system of the electric utilities is assuming an even greater problem since the aggravation of climate change and extreme weather. The vegetation growth cycle and the stress associated with the more basic infrastructure are also impacted by regulations and worsen the likelihood of local outages due to the higher temperatures, change in the storm patterns, and prolonged droughts (Ramli, 2023) [28]. Similarly, in most areas' distribution assets are approaching or even surpassing their design life, making them susceptible to equipment-related failure. All this has changed the feeder-level outage prediction not only into a technical problem but also into a regulatory and a policy problem since utilities are now accountable for resilience planning and reliability requirements in the context of climate-related risk (Zhu, 2025) [40].

Though time-based outage risk models have long been utilized, they are generally not successful when exposed to very nonlinear drivers of outage such as vegetation intrusion or weather-driven compounds. Classical models that have been widely utilized and applied to forecasting in the short horizon include ARIMA, ETS, and Prophet because they are interpretable and are based on well-established statistical frameworks (Ali, 2025; Suryawan, 2024) [2, 37]. However, these models often struggle with capturing complex, nonlinear patterns in data, prompting the exploration of hybrid approaches that integrate machine learning techniques to enhance forecasting accuracy (Arumugam, 2025; Ezugwu, 2025) [4, 10]. Classical statistical models, such as ARIMA, are grounded on statistics (Ali, 2025) [2]. They are, however, fairly stationery and experience issues with capturing interactions between features, particularly where the outage pattern is also determined by the weather conditions, vegetation indices, and the combination of equipment age (Sabzevari, 2025) [32]. Recent developments in hybrid statistical models have attempted to address these shortcomings, but they have not yet been scaled to many different feeders (Ikrimat, 2025) [14].

Concurrent advances in machine learning have expanded the methodological repertoire of utilities. Trees and recurrent neural network ensemble models, in particular LSTMs, have been shown to be more accurate in modeling outage risk in the case of nonlinear dynamics (Kotsompolis, 2025; Ning, 2022) [17, 24]. However, their real-world application has been pretty inconsistent, partly because of explainability issues and the hard task of integrating model knowledge into maintenance protocols (Gulyaev, 2025; Rowan, 2025) [11, 31]. In utilities, one really needs to balance predictive performance and interpretability because actionable forecasts have to meet the expectations of the regulators and be useful for decision-making at the field level (Aror, 2025) [3].

This research actually bridges those gaps by comparing the classic time series models systematically with tree-based and deep learning models on short-horizon outage prediction at the feeder level. Besides being accurate in predicting the likelihood of an outage, the study's authors added explainability via SHAP-based feature attribution, which enables utilities to see how, for example, weather, vegetation, or equipment factors drive outage risk (Ng, 2025; Raykar, 2023) [23, 29]. This integration is super helpful for the vegetation management plan and deciding on equipment replacement priorities, and the predictions tie straight into operational strategy (Noviandy, 2025) [26]. Additionally, it feeds into the current discussion of explainable AI in infrastructure, offering a good balance between predictive maintenance and matching reliability with forecasting (Bachmann, 2025) [5].

Objectives and Contributions

- Benchmark ARIMA, ETS, and Prophet against tree-based ensembles and LSTM models for feeder-level outage forecasting (Ali, 2025; Suryawan, 2024) [2, 37].
- Implement rolling-origin evaluation, calibration, and reliability diagrams to assess forecast robustness (Lavanya, 2024) [19].
- Apply SHAP-based feature attribution to interpret model outputs and guide vegetation management and maintenance prioritization (Ng, 2025; Noviandy, 2025) [23, 26].

2. Literature Review

2.1. Outage Forecasting in Power Systems

Forecasting outage events has long been a priority for utilities seeking to improve resilience and service continuity. Machine learning is increasingly becoming popular in recent years as utilities are becoming responsible for both regulatory performance and customer service measures (Zhu, 2025) [40]. Majority of the studies are centered on measurements derived out of weather reports, vegetation indexes, or equipment databanks to forecast outage likelihood at either the regional or system-wide levels. These larger studies demonstrate that additional nontraditional predictors such as radar-tracked storms may be useful, but they may otherwise fail to predict at the fine-grained feeder level where outages do occur (Sabzevari, 2025) [32]. The combination of outage prediction and load or demand representation has made it open to new approaches, particularly when making environmental variables dependent on consumption patterns. As an example, studies on hybrid models show that outage prediction, which is based on weather and combined with load forecasting, may be applied to enhance the short-term planning and asset management (Soumbara, 2025; Ezugwu, 2025) [35, 10].

2.2. Statistical Models of Time Series

The classical time-series approaches have dominated the dominant role in outage risk/reliability modelling, as they are mathematically transparent and have historical experience in the engineering venture. One such option is the ARIMA model, which is relevant to a range of other disciplines, including engineering or economics, since it can be used for time-dependent autocorrelations of outage or demand data (Ali, 2025; Ikrimat, 2025) [2, 14]. It has been helpful to model seasonality in reliability work and the persistence of events over time yet remains with a nonlinear outage driver problem (Sabzevari, 2025) [32]. Newer changes strive to make it flexible, though the computational cost of frequency updates can make it difficult to execute in real time on utility systems (Ikrimat, 2025) [14].

Exponential smoothing (ETS) models can also help eliminate the propensity to overlook the trend and seasonality (Ali, 2025; Casonatto, 2025) [2, 6], which is another set of tools that is time-tested and often used in export and financial forecasts. Their primitive parameterization made them appealing in reliability predictions where utilities sought sparse real-time models. According to macroeconomic research, ETS is also commended for providing consistent projections even under swampy situations (Casonatto, 2025) [6]. Nonetheless, ETS is not able to detect the instantaneous outage bursts that occur when there are compound weather and vegetation events (Sabzevari, 2025) [32].

Prophet has gained more popularity in recent times, as it automatically handles seasonality and can automatically extract domain knowledge through custom regressors. It began in demographics and finance and is now applied in areas that require transparency and fast installment (Narwal, 2024; Hakkar, 2025) [22, 12]. In practice, where it is praised due to its ability to be adapted and scaled, it still fails to perform consistently with very volatile data on outages (Hakkar, 2025) [12]. Overall, ARIMA, ETS, and Prophet remain useful reference standards of reliability forecasts but cannot capture the vegetation-weather interaction or unexpected nonlinear equipment breakdowns due to linear assumptions (Suryawan, 2024) [37].

2.3. Machine Learning Models

The limitations of traditional statistical methods have prompted utilities and researchers to explore machine learning models that can capture nonlinear dependencies and higher-order feature interactions. Gradient-boosted trees are the new hype of the day, especially when used in conjunction with hybrid architectures that are good at predicting outage risks because they are able to combine clean and structured operational data with unstructured and messy environmental data (Kotsompolis, 2025; Liu, 2025) [17, 20]. The advantage of the matter is that they are highly effective in ranking predictive variables, and thus, they assist in labeling maintenance and vegetation issues, which demand the greatest consideration. Nevertheless, it is not a tree only but a collection in the air where the residual learners are stacked to push the predictions even further (Santos, 2023) [34].

The recurrent neural networks and LSTMs, specifically, have become some of the most popular in tracking time-dependently the outage and load data anomalies. Nonlinear seasonality Folks in the energy space have reported a more successful direction of learning the nonlinear seasonality, which sheds solid projections of power price and weekly demand projections (Ning, 2022; Kandpal, 2023) [24, 15]. Predicting outages internally by predicting using weather indices and vegetation indices has proved to be intriguing, advancing the risk assessment in a fast scan at the feeder end (Suryawan, 2024) [37]. Hybrid models involving traditional statistics as well as the application of modern machine learning are also emerging. One such example is that ARIMA-ANN, or ARIMA-ML, is trying to retain the interpretability of the standard ARIMA and permits an ANN to learn the nonlinear residuals (Arumugam, 2025; Elseidi, 2024; Diaz, 2025) [4, 9, 8]. ARIMA hybrids are also ARIMA hybrids that attempt to be more stable in a variety of conditions (Elseidi, 2024) [9]. In a similar manner, the residual-based ensemble trees are also good in outage prediction, a factor that mitigates the bias that would otherwise be observed when using a single model (Santos, 2023) [34]. These ML tricks, however, are linked with trade-offs: more accuracy can result in overfitting or higher compute, and a real-world problem of scaling to live utility processes (Gulyaev, 2025) [11].

2.4. Utility Analytics Explainability

Predictive accuracy is one of the aspects of concern, but utilities should make sure that their outage prediction model can be readable to enable the effective pace of the risk management practices and adherence to the check. Explicit machine learning is also desirable, especially in the controlled setting wherein the operators will have to be prepared to justify their decisions concerning the maintenance schedule and vegetation management (Aror, 2025) [3]. As engineering research demonstrates, so-called black-box models, even with super accurate models, are pushed back against until they provide some insight into what is driving risk (Rowan, 2025; Tatsat, 2025) [31, 38].

SHAP-based feature attribution is now the default choice when it comes to explainability, as the predictions of a model can be traced to important input drivers. New developments have extended SHAP to structured and time-series data, providing both local and global interpretability of results (Ng, 2025; Raykar, 2023) [23, 29]. Temporal extensions such as TsSHAP have been constructed with temporal sequences in mind, and they assist in explaining recurrent neural network predictions and reveal the effect of weather or vegetation patterns on the risk of outage (Raykar, 2023) [29]. SHAP has been used successfully in energy predictive control to wrap an ensemble tree model and allow utilities to visualize the influence of each weather parameter on their outage risk over time (Noviandy, 2025) [26]. With all these developments, the question of the limit of interpretability and the trade-off between transparency and complexity is still debated among people. Some researchers claim that SHAP is a highly practical tool that can produce an artificial sense of causality in high-dimensional space (Righetti, 2022) [30]. Other people also point out that they have to tradeoff between explainability and predictive power, especially when the operational safety of the decision is based on the accurate predictions (Tatsat, 2025) [38]. In the utility case, such discussions highlight the need to adopt explainable forecasting models that will deliver sound performance and operating outcomes that can be utilized in the decision-making process in maintenance (Bachmann, 2025) [5].

3. Data Sources & Preprocessing

It is just a matter of making predictive forecasts on past outage records at the feeder level; nevertheless, it is a pain to prepare such data to be studied. All utilities have outage tickets in their format, and thus cannot in any manner start to model them without having to reconcile them. The implication of that in practice is that the ticket timestamps would be matched with crew logs, and it is guaranteed that all the events are allocated to the appropriate feeder and not confused with the other service events. In practice the metadata is more likely to be disparate, or pieces of the records are absent, and this is a common bane of utilities. Cleaning and formatting all this is therefore a significant step towards the production of reliable predictive properties (Zhu, 2025) [40]. The same preprocessing makes between-system duplicates and outage event code clears much easier so that the outage cause categories fit modeling dreams.

Another massive piece of the puzzle of outage forecasting is weather and radar information. The variability in climate is gradually gaining more and more power over the dependability of distribution, and this is the reason why weather predictors should be introduced. One such area is wind forecasting, where it is now combined with physics-based models as well as machine learning in order to improve the short-term accuracy (Alam, 2025; Elseidi, 2024) [1, 9]. These approaches indicate the applicability of the fine-resolution weather variables to the outage prediction. The risk of outage has been shown to be strongly correlated with radar-based precipitation and storm-intensity scales, as well as multi-scale weather data, which has proven useful in both the local effects of storms and the macro-effects of the season (Sabzevari, 2025) [32]. Incorporating radar measurements and ground-station measurements, it is possible to gain a better understanding of environment-related stressors that affect the quality of feeder services, and the models will be more applicable to other areas (Elseidi, 2024) [9].

Additionally, the vegetation is the other predictor that is of a specific significance in which vegetation is directed to bad weather and pulls the line faults up. Large-scale density and health of vegetation can be measured using satellite-based vegetation indices like normalized difference vegetation index (NDVI). The results of NDVI aggregation will enable to track the cycle of vegetation development and potential interference with the canopies (Zhu, 2025) [40]. Not only does the NDVI make the models more accurate but also operational planning since the vegetation-management schedules may be linked to these quantitative indicators. Nevertheless, there remains the need to preprocess satellite properties carefully, such as cloud cover, spatial resolution, and time lags in data availability, so that the plant cover signals would be properly synchronized with the outage schedule (Sabzevari, 2025) [32].

The asset-related characteristics, together with the weather conditions and vegetation, are significant in the representation of the feeder-specific weaknesses. The age of equipment, such as that of transformers, insulators, and breakers, has been repeatedly determined to be an important factor in outage frequency whereby old equipment like this would be discarded to fail in agitating conditions (Soumbara, 2025) [35]. The integration of the asset registry information and the outage information should be well bonded because the difference in labeling and poor maintenance records have the disposition of making this a complicated process. The crew logs can also be used as a supplement to such datasets to give specifics about the response time, the complexity of the repair, and the most frequent points of failure, and thus put the reality of operations in a modeling system (Zhu, 2025) [40].

The last phase in the data preparation process is feature engineering, in which the raw variables are engineered into predictive data, which can be fed into both time-series and machine learning models. Rolling-origin assessment has become an efficient utility forecasting model to ensure that the training and evaluation window can allow the time series of outage events to be considered and can provide solid estimates of the stability of the model (Lavanya, 2024) [19]. The feature selection art is also significant since in cases where a model consists of a set of redundant or noisy predictors, the model will be unstable or is likely to inflate the variance. The latest studies have taken notice of dimensionality reduction methods, which are interpretable and computationally efficient and can be integrated with large-scale weather and vegetation data without overfitting (Pabuccu, 2024; Noviandy, 2025) [27, 26]. They can also be more efficiently selected using more efficient feature selection methods like SHAP-informed pruning to enable modelers to only select the most important predictors to achieve higher forecasting accuracy and transparency (Noviandy, 2025; Bachmann, 2025) [26, 5]. The result of all these preprocessing activities is the production of high-quality structured data, which can be used in the comparative analysis of the machine learning and the statistical approaches to outage risk prediction.

4. Methodology

The situation in the current study is the short-horizon prediction of the risk of feeder-level outage whereby the objective is to forecast the potential failures between the daily and weekly levels. This scope has a suitable balance between the feasibility of actionable forecasts and the feasibility of modeling localized outages on the basis of data. Short-horizon predictions in distribution feeders are also highly applicable in regions where there is an overlap of vegetation growth, weather shocks, and deterioration of assets to combine and pose reliability threats (Ramli, 2023; Zhu, 2025) [28, 40]. The reduction of the horizon to the feeder level allows the utilities to align the predictive analytics to the maintenance scheduling and trimming models and the emergency crew location, therefore causing the model results to be directly converted into the preventive actions (Matenga, 2025) [21].

The model system has four forecasting techniques classes. The former includes statistical techniques such as ARIMA, ETS, and Prophet, which are long-established in the engineering, economics, and financial sectors. These models have proved very effective in situations where the data is arranged and seasonal, although they lack the capabilities of modeling non-linear data of outages since they are founded on linear assumptions (Ali, 2025; Narwal, 2024) [2, 22]. ARIMA and ETS at least give a reference point on evaluation of forecast performance (Ali, 2025) [2], whereas Prophet is already an accessible tool in the process of complex elements of seasonality and trend in financial and demographic data (Hakkar, 2025) [12]. The second category of algorithms is tree-based algorithms, that is, gradient boosting and XGBoost, which have been discovered to be more precise in regard to the ability to capture non-linear relationships and high-order interactions. They are very useful in outage risk applications since they have heterogeneous features like weather, vegetation indices, and asset characteristics (Kotsompolis, 2025; Liu, 2025) [17, 20].

The third type also is deep learning models, and it is focused on long short-term memory (LSTM) networks. The latter architectures come in particularly handy in the modelling of sequential dependencies of time series, so that they can model quite complex lag structures not captured by traditional ones. The outcomes of LSTM-based hybrids are already high regarding the energy demand forecasting and oil production experiments, which proves that they can be applied to the concept of reliability (Ning, 2022; Suryawan, 2024) [24, 37]. The present trends further broaden their ability to utilize hybrid structures of LSTMs combined with decomposition operation or attention mechanism to allow multi-scale outage driver representation (Gulyaev, 2025; Liu, 2025) [11, 20]. Additionally, a hybrid between statistical ML and some in-between models combines the interpretability of ARIMA with the power of artificial neural nets or ensemble learning. An example of this is the use of ARIMA-ANNs to produce success in load prediction (Ezugwu, 2025; Arumugam, 2025) [10, 4], and retail and energy systems have been demonstrated to be improved by hybrid frameworks that incorporate classical residual modeling and machine learning advancements (Choubey, 2025; Santos, 2023) [7, 34]. This type of diversity of models makes sure that the contrast between the old statistical techniques and the new ones provided by machine learning is made.

The model is evaluated with the systematic design, which focuses on the predictive accuracy and reliability. The most significant feature of the experimental model is the rolling-origin evaluation that provides a time validity of the model training by the use of the past window and predicts the future. The approach offers a more realistic evaluation of how models behave under realistic conditions and when the future data is non-existent (Lavanya, 2024) [19]. A combination of both point forecast performance and probabilistic calibration is measured using measures of forecast performance, such as mean absolute percentage error, root mean squared error, and continuous ranked probability scores (Nortey, 2025) [25]. The reliability of predictive outage forecasting models to capture the distribution of risks of a feeder also needs to be assessed on the basis of forecasting accuracy. Reliability diagrams and calibration plots are employed to determine how closely the predicted probabilities can replicate the observed outage frequencies in such a way that the utilities can be in a position to trust the forecasts as well as the uncertainty margins of the same forecasts (Ramli, 2023) [28].

An explainability layer is adopted as a supplementary layer to accuracy reviews through SHAP-based feature attribution. The scheme allows a universal and local understanding of how the weather, vegetation, and asset-related effects on the reported outage risk can be identified to assist utilities in concentrating on the most significant drivers (Ng, 2025; Noviandy, 2025) [23, 26]. The efficiency concern is taken into account with optimal SHAP approximations at the expense of a reduced computational overhead and the lack of insight in terms of selling (Bachmann, 2025) [5]. Moreover, it is contrasted with time-series-specific models such as TsSHAP, which offer great model-agnostic interpretability to the time-varying dynamics (Raykar, 2023) [29]. All these methods render the findings of the forecasting process to be credible, transparent, and operationally efficient.

5. Results

5.1. Forecast Accuracy Comparison

The comparative benchmarking indicated that hybrid models had always been better than classical statistical tools in accuracy and robustness. Hybrid ARIMA-ANN setups generated much smaller forecast errors compared to the traditional ARIMA or exponential smoothing forecasts, which is their capacity to forecast linear and nonlinear outage dynamics (Arumugam, 2025; Ezugwu, 2025) [4, 10]. Similarly, gradient boosting trees and LSTM networks demonstrated significant differences in performance in terms of their capability to model intricate temporal interplay and particularly in the case of dissimilar weather-associated risks (Soumbara, 2025; Kotsompolis, 2025) [35, 17].

The cases were being directly compared, which showed differences in performance between families of models. ARIMA was interpretable and valid when the outage risk profiles did not change radically over time because of steady longer-term trends but failed to respond to outage risk profiles that changed quickly (Ali, 2025) [2]. Prophet also was better adaptable to seasonal variations and less precise at the feeder-level granularity (Yadav, 2023) [39]. Compared to both statistical counterparts, LSTM networks could learn to develop nonlinear understanding of sequence-based relationships and are more robust to changes in regimes in weather and vegetation conditions (Sunki, 2024; Suryawan, 2024) [36, 37]. Another finding of Suryawan *et al*. (2024) [37] is that the long-memory properties of LSTM could conduct a more excellent generalization, especially in high-variability areas [37]. Overall, these results demonstrate that machine learning and hybrid solutions to short-term feeder-level risk forecasting are beneficial.

5.2. Calibration & Reliability

In addition to the accuracy, the reliability of models became one of the determinants to practice. The reliability diagrams showed that the well-calibrated hybrid and the deep learning models better matched the observed frequencies of outages than did the statistically based approaches (Ramli, 2023) [28]. The models that had a high predictive variance were prone to overconfidence, particularly when the situations of outage were rare but large and the calibration error was overstated.

Essential trade-offs were demonstrated by failure to plan across the different forecasting horizons. At very short horizons, models based on trees and models based on LSTM were found to be more credible, but their calibration again needed to be done cautiously to avoid over- or under-predicting the occurrence of rare spikes in the outage likelihood (Kotsompolis, 2025) [17]. The statistical models like ETS were found to be more accurate at longer horizons but less accurate (Kouokam, 2025) [18]. Hybrid ARIMA-ANN techniques were a tradeoff of these weaknesses, being flexible in the short run and capturing long-term trends, and providing the minimum total calibration errors (Ezugwu, 2025) [10]. These are the major points that calibration does not just happen to be a statistical necessity but a determinant of whether the predictions can be translated into working operational planning.

5.3. Explainability Insights

Integration of SHAP was able to give information on the major agents of outage hazards. The global features rankings continually showed that the best predictors were weather phenomena, particularly wind velocity and quantity, and vegetation growth indicators that were calculated via the NDVI indices (Zhu, 2025; Sabzevari, 2025) [40, 32]. Local attributions were used to support the fact that abrupt alterations in vegetation density enhanced the probability of outage during even moderate weather conditions, which also supported the interaction effects that are usually dropped in univariate models (Noviandy, 2025) [26].

Additional interpretation was achieved through dimensionality-conscious explainability techniques. Similar to high-dimensional spaces of Noviandy *et al*. (2025) [26], SHAP is healthier in those scenarios when the feature space complexity is reduced, and predictors with large values are selected and eliminate noise in ascribing variables [26]. The solution facilitated mapping of the sources of operational data and outage risk prediction in a better way and made available to the utility stakeholders interpretable and actionable insights. This literature has shown in general that explainability does not merely render the behavior of models justifiable but also assists in reducing the gap between data-driven forecasts and regulation or operational accountability (Bachmann, 2025) [5].

6. Discussion

6.1. Interpretation of Findings

The results of the current study are a part of the constantly growing literature, indicating that hybrid techniques are superior to classical outage forecasting techniques. Just as Casonatto *et al*. (2025) [6] found that statistical and machine learning models are more robust when combined together in the event of dynamic conditions of the system, the same was witnessed in the present study [6]. Similarly, Choubey (2025) [7] highlighted the fact that ARIMA-ANN hybrids are the most robust in order to balance interpretability and predictive power, which are most aligned with the listed advantages [7]. Similar to the shown calibration enhancement of the hybrid and deep learning models, Nortey *et al*. (2025) [25] also pointed to the fact that the accuracy alone is insufficient, but the probability values must be well-calibrated [25].

Among the results that can be critically observed due to benchmarking, the specific advantage of nonlinear use of hybrid and machine learning is singled out. Han (2025) [13] claims that classical models suffer problems with structure breaks and non-standard patterns, and the results of this paper prove that the outages of vegetation bursts and abnormal weather cannot be adequately explained with the help of the extrapolation of the linear models [13]. Diaz *et al*. (2025) [8] have also shown that the effective model of nonlinear dependencies would not only be more accurate but also more flexible in highly uncertain cases [8]. The complementary supporting findings contribute to achieving the necessity to change the paradigm of the traditional model of statistical dependence to the hybrid and deep learning model as a classical outage predicting device.

6.2. Maintenance Prioritization Implications

The proposed layer of explainability by SHAP has a direct implication of maintenance prioritization in utility networks. SHAP allows operators to define those features (wind intensity, precipitation, and vegetation indices) that put the largest number of stressors on the feeders, putting preemptive maintenance schedules first (Ng, 2025; Noviandy, 2025) [23, 26]. This operationalizes the outputs of the forecasting to the tangible strategy of setting up resources to reduce the cost of the outage and taking less time to restore the company. The relevance of vegetation indices to the development of the outage risk demonstrated by Zhu and Quiring (2025) [40] also proves the importance of including the ecological monitoring in the practices of outage management in the assets [40].

In a broader risk management perspective, AI tools can be used to augment the decision-making process by adding some form of predictive intelligence to the maintenance planning process. Aror and Mupa (2025) [3] believe that AI solutions are gaining critical importance in identifying high-impact risks and also making sure that a few resources are spent on high-priority interventions due to their limited budget [3]. These abilities of hybrid models to learn insights into the nonlinear triggers along with interpretability using SHAP provide a predictive maintenance model that is not only technically sound but also economically viable in the case in point. This reactive-to-anticipatory maintenance shift is an instance of the paradigm shift in utility risk management that is enabled by the advanced methods of utility forecasting and explicability.

6.3. Integration into Utility Workflow

The operational innovation is applicable because the introduction of predictive and explainable forecasting tools as utility processes has been introduced. Matenga *et al.* (2025) [21] also identified similarities in the energy efficiency and mechatronics, where the predictive analytics have already been translated into the ability to maximize the scheduling and proactive system tuning [21]. The utility cases can be conducted according to the same principles; the outage prediction is used in both cases to regulate the work of the grid and the asset life cycle as well. The tools are scalable to the system-level planning and not only to the localized feeder-level analysis because of their scalability (Soumbara, 2025) [35].

Also, the user psychology dashboards and decision support systems need to be prepared to inculcate the knowledge of forecasting into daily practice. The visualization tools must be of quality, as Rowan and Doostan (2025) [31] noted, because the complex model outputs must be linked to the practitioner decision-makers [31]. In that regard, incorporation of SHAP elucidation into a live dashboard will create transparency in a manner that the operators may have the chance to track the expected dangers to the particular drivers (Ng, 2025) [23]. This form of integration is not only effective in terms of operations but also improves regulatory compliance, as they are able to use auditable evidence to justify risk-informed decisions (Aror, 2025) [3].

7. Limitations

This study also has some limitations, in particular, the quality and heterogeneity of data sources. Crew logs and outage tickets are not necessarily full as well, and they are not always coded; this may introduce certain bias into the model training (Ramli, 2023; Zhu, 2025) [28, 40]. Similarly, heterogeneous sources (radar data, vegetation indices, and asset records) require critical preprocessing and thus may cause reduced reproducibility due to the format and time scale harmonization (Santos, 2023) [34].

The other weakness is that deep learning models, especially the LSTMs, can be interpreted. Despite a certain amount of transparency of SHAP, the given nature of recurrent networks is inclined to create a veil over the specific mechanisms that are used to differentiate how forecasts are formed, and there is a challenge of finding anyone responsible at the operational level (Tatsat, 2025; Rowan, 2025) [38, 31].

Additionally, geographic generalizability of the findings may be limited. According to Soumbara *et al*. (2025) [35], outage drivers are significantly varied depending on climatic conditions and regulatory circumstances [35]. The models that were optimized with the data of one of the regions may not be directly applicable to the regions that have the dissimilar vegetation profile or types of infrastructure. The framework must therefore be demonstrated on a wider scope in other geographies before it can be proposed that it should be adopted on a wide scale (Sabzevari, 2025) [32].

8. Conclusion & Future Work

The paper has an application to the utility outage forecasting field since the paper shows a demonstration of how a hybrid pipeline may be utilized to forecast power outages and failures involving a statistical and machine learning model with an explainability layer. The study that compares ARIMA, ETS, and Prophet with tree-based ensembles and LSTM hybrids indicates the disadvantages and advantages of other forecasting families (Ali, 2025; Suryawan, 2024) [2, 37]. SHAP, which is built into the model as a post hoc interpretability tool, ensures that the model predictions can be transformed into actions to be implemented on the feeder level for maintenance and vegetation management (Ng, 2025; Bachmann, 2025) [23, 5].

Significant findings are that hybrid models are never less precise compared to classical models of time series, particularly when there are nonlinear interactions such as vegetation growth and extreme weather (Choubey, 2025; Diaz, 2025) [7, 8]. However, reliability is not sufficient to be introduced operationally. Explainability introduced with the help of SHAP and TsSHAP methods enhances trust and transparency technologies according to which the operators are in a position to adjust predictive outputs in line with the regulation compliance and risk management systems (Raykar, 2023; Noviandy, 2025) [29, 26].

There are several avenues that one should explore as a prospect. The additional inputs provided by the smart grid technologies and IoT-based sensors to the model may be finer and real-time and, consequently, responsive (Matenga, 2025) [21]. The SHAP causation methods also offer the possibility of moving the correlation-based explanations towards causal inferences to render them more interpretable (Ng, 2025) [23]. Hybrid optimization methods also give the opportunity for further enhancement of predictive reliability, such as dynamic weighting of the statistical and ML factors (Liu, 2025; Arumugam, 2025) [20, 4]. Combined, these guidelines lead to the fact that hybrid models require constant improvements in order to become more precise, besides providing transparency and scalability of their operations.

References

1. Alam U, Mozomder AS. Comparative machine learning and time series forecasting of wind power output using SCADA data. J Comput Sci Technol Stud [Internet]. 2025 [cited 2025 Oct 10];7(7):14-30. doi:10.32996/jcsts.2025.7.7.2
2. Ali MP, Sadia AZ, Khanam M. ARIMA vs. ETS for RMG export forecasting of Bangladesh: a comparative study on model accuracy. Dhaka Univ J Sci. 2025;73(2):101-5. doi:10.3329/dujs.v73i2.82767
3. Aror TA, Mupa MN. Risk and compliance paper what role does Artificial Intelligence (AI) play in enhancing risk management practices in corporations? World J Adv Res Rev. 2025;27(1):1072-80. doi:10.30574/wjarr.2025.27.1.2607
4. Arumugam V, Natarajan V. Enhanced time series forecasting using hybrid ARIMA and machine learning models. Indones J Electr Eng Comput Sci. 2025;38(3):1970-9. doi:10.11591/ijeecs.v38.i3.pp1970-1979
5. Bachmann S. Efficient XAI: a low-cost data reduction approach to SHAP interpretability. J Artif Intell Res. 2025;83. doi:10.1613/jair.1.18325
6. Casonatto R, da Silva ECM, Barbalho S, Gonçalves M, Gabriela M. Comparative analysis of demand forecasting models: a time series case study [Internet]. 2025 [cited 2025 Oct 10]. doi:10.20944/preprints202509.1013.v1
7. Choubey A. Hybrid models for retail demand forecasting: integrating classical time-series and machine learning approaches. Asian J Res Comput Sci [Internet]. 2025 [cited 2025 Oct 10];18(8):113-23. doi:10.9734/ajrcos/2025/v18i8744
8. Diaz JM, Ahmad A, Arshad M. Time series analysis: hybrid econometric-machine learning model for improved financial forecasting. J Comput Biomed Inform [Internet]. 2025 [cited 2025 Oct 10];9(2). doi:10.56979/902/2025
9. Elseidi M. MSTL-NNAR: a new hybrid model of machine learning and time series decomposition for wind speed forecasting. Stoch Environ Res Risk Assess [Internet]. 2024 [cited 2025 Oct 10]. doi:10.1007/s00477-024-02701-7
10. Ezugwu CO, Opata CE, Dimeze AI, Ogbuka CU, Nwosu CM. Development of a hybrid ARIMA-ANN model for enhanced electric load forecasting accuracy [Internet]. 2025 [cited 2025 Oct 10]. Available from: https://www.researchgate.net/publication/390640827_DEVELOPMENT_OF_A_HYBRID_ARIMA-ANN_MODEL_FOR_ENHANCED_ELECTRIC_LOAD_FORECASTING_ACCURACY
11. Gulyaev AV. Hybrid deep learning model for time series forecasting. Neurocomputers [Internet]. 2025 [cited 2025 Oct 10];(3):62-72. doi:10.18127/j19998554-202503-08
12. Hakkar F. Intelligent forecasting of GBP/USD exchange rates: a comparative analysis of ARIMA, Prophet, and LSTM models [Internet]. 2025 [cited 2025 Oct 10]. doi:10.13140/RG.2.2.16561.29286
13. Han S. Financial time series forecasting: a hybrid approach combining AR-GARCH and machine learning models. Trans Comput Sci Intell Syst Res. 2025;10:72-7. doi:10.62051/pg9aec47
14. Ikrimat A, Akobi C, Peter PT, Gatta AA. Forecasting Nigeria inflation rate using Autoregressive Integrated Moving Average (ARIMA) model. Mikailalsys J Adv Eng Int [Internet]. 2025 [cited 2025 Oct 10];2(2):201-17. doi:10.58578/mjaei.v2i2.5649
15. Kandpal PK, Shourya, Yadav Y, Sharma N. Time series forecasting of NSE stocks using machine learning models (ARIMA, Facebook Prophet, and Stacked LSTM). Lect Notes Netw Syst. 2023;303-20. doi:10.1007/978-981-99-6553-3_24
16. Kim SY, Du Y, Zhang T, Shen J, Hou Y. Adolescent executive function as a resilience factor in the family stress model among Mexican-origin families. J Fam Psychol [Internet]. 2025 [cited 2025 Oct 10]. doi:10.1037/fam0001386
17. Kotsompolis G, Cheilas P, Konstantakis KN, Sfakianakis E, Michaelides P. Smart forecasting of carbon prices using machine learning and neural networks: when ARIMA meets XGBoost and LSTM. J Forecast [Internet]. 2025 [cited 2025 Oct 10]. doi:10.1002/for.70025
18. Kouokam LS, Pemberton J, Rodrigues PC. Hybrid time series forecasting models: an in-depth evaluation across financial and climatic indicators. Fluct Noise Lett. 2025. doi:10.1142/s0219477525500245
19. Lavanya H, Pratyusha P. A comparative study of rolling versus benchmark performance returns in selected mutual funds. Int J Sci Res Eng Manag. 2024;08(09):1-15. doi:10.55041/ijsrem37576
20. Liu Z, Zhang Z, Zhang W. A hybrid framework integrating traditional models and deep learning for multi-scale time series forecasting. Entropy [Internet]. 2025 [cited 2025 Oct 10];27(7):695. doi:10.3390/e27070695
21. Matenga HR, Mupa MN, Musemwa OB. Artificial intelligence-driven mechatronic system for energy efficiency in U.S manufacturing industries [Internet]. 2025 [cited 2025 Oct 10];9(3):792-801. Available from: https://www.researchgate.net/publication/395573582_Artificial_Intelligence-Driven_Mechatronic_System_for_Energy_Efficiency_in_US_Manufacturing_Industries
22. Narwal R, Aggarwal H. ARIMA, Prophet, and LSTM-based analysis of demographic factors in smartphone usage patterns. Microsyst Technol [Internet]. 2024 [cited 2025 Oct 10]. doi:10.1007/s00542-024-05734-8
23. Ng WY, Wang LR, Liu S, Fan X. Causal SHAP: feature attribution with dependency awareness through causal discovery [Internet]. 2025 [cited 2025 Oct 10]. doi:10.48550/arXiv.2509.00846
24. Ning Y, Kazemi H, Tahmasebi P. A comparative machine learning study for time series oil production forecasting: ARIMA, LSTM, and Prophet. Comput Geosci. 2022;164:105126. doi:10.1016/j.cageo.2022.105126
25. Nortey ENN, Agyemang EF, Sakyi-Yeboah E, Ampomah O-A, Agyekum L. AI meets economics: can deep learning surpass machine learning and traditional statistical models in inflation time series forecasting? Data Sci Finance Econ [Internet]. 2025 [cited 2025 Oct 10];5(2):136-55. doi:10.3934/DSFE.2025007
26. Noviandy TR, Idroes GM, Lala A, Helwani Z, Idroes R. Inductive biases in feature reduction for QSAR: SHAP vs. autoencoders. Infolitika J Data Sci [Internet]. 2025 [cited 2025 Oct 10];3(1):40-9.

doi:10.60084/ijds.v3i1.306

27. Pabuccu H, Barbu A. Feature selection with annealing for forecasting financial time series. Financ Innov. 2024;10(1). doi:10.1186/s40854-024-00617-3
28. Ramli N, Ab M, Mohamad D, Othman M, Nassir AA. Fuzzy time series forecasting accuracy based on hybrid similarity measure. J Sci Math Lett [Internet]. 2023 [cited 2025 Oct 10];11(2):93-103. doi:10.37134/jsml.vol11.2.11.2023
29. Raykar VC, Jati A, Mukherjee S, Aggarwal N, Sarpatwar K, Ganapavarapu G, *et al*. TsSHAP: robust model agnostic feature-based explainability for time series forecasting [Internet]. 2023 [cited 2025 Oct 10]. doi:10.48550/arxiv.2303.12316
30. Righetti G, Porello D, Confalonieri R. Evaluating the interpretability of threshold operators. Lect Notes Comput Sci. 2022;136-51. doi:10.1007/978-3-031-17105-5_10
31. Rowan C, Doostan A. On the definition and importance of interpretability in scientific machine learning [Internet]. 2025 [cited 2025 Oct 10]. doi:10.48550/arXiv.2505.13510
32. Sabzevari Y, Eslamian S, Sleziak P, Okhravi S. Hydro-climatic time series analysis of wind speed in western Iran: an ARIMA approach [Internet]. 2025 [cited 2025 Oct 10]. Available from: https://www.researchgate.net/publication/393899060_Hydro-climatic_time_series_analysis_of_wind_speed_in_western_Iran_an_ARIMA_approach
33. Samuel D, Mohanraj V, Velayutham S, Janakiraman S, Vijayagopal V. Improving time series forecasting accuracy with hybrid radial basis function networks. Adv Comput Intell Robot. 2024;419-36. doi:10.4018/979-8-3693-4326-5.ch019
34. Santos D, Neto PSGD, de Oliveira JFL, Cavalcanti GDC. A hybrid system based on ensemble learning to model residuals for time series forecasting. Inf Sci. 2023;649:119614. doi:10.1016/j.ins.2023.119614
35. Soumbara S, Moulim A, Ghini AE. Hybrid time series and neural network model for forecasting electricity demand under climate constraints. J Model Manag. 2025. doi:10.1108/jm2-05-2025-0199
36. Sunki A, SatyaKumar C, Narayana GS, Koppera V, Hakeem M. Time series forecasting of stock market using ARIMA, LSTM and FB prophet. MATEC Web Conf [Internet]. 2024 [cited 2025 Oct 10];392:01163. doi:10.1051/matecconf/202439201163
37. Suryawan IGT, Putra IKN, Meliana PM, Sudipa IGI. Performance comparison of ARIMA, LSTM, and Prophet methods in sales forecasting. Sinkron J Penelit Tek Inform [Internet]. 2024 [cited 2025 Oct 10];8(4):2410-21. doi:10.33395/sinkron.v8i4.14057
38. Tatsat H, Shater A. Beyond the black box: interpretability of LLMs in finance [Internet]. 2025 [cited 2025 Oct 10]. doi:10.48550/arXiv.2505.24650
39. Yadav S. A comparative study of ARIMA, Prophet and LSTM for time series prediction. J Artif Intell Mach Learn Data Sci. 2023;1(1):1813-6. doi:10.51219/jaimld/sandeep-yadav/402
40. Zhu L, Quiring SM. Hybrid machine learning model for hurricane power outage estimation from satellite night light data. Remote Sens [Internet]. 2025 [cited 2025 Oct 10];17(14):2347. doi:10.3390/rs17142347